

Compressed Sensing with coherent tight frames via l_q -minimization for $0 < q \leq 1$ *

Song Li and Junhong Lin[†]

Department of Mathematics, Zhejiang University
Hangzhou, 310027, P. R. China

Abstract

Our aim of this article is to reconstruct a signal from undersampled data in the situation that the signal is sparse in terms of a tight frame. We present a condition, which is independent of the coherence of the tight frame, to guarantee accurate recovery of signals which are sparse in the tight frame, from undersampled data with minimal l_1 -norm of transform coefficients. This improves the result in [1]. Also, the l_q -minimization ($0 < q < 1$) approaches are introduced. We show that under a suitable condition, there exists a value $q_0 \in (0, 1]$ such that for any $q \in (0, q_0)$, each solution of the l_q -minimization is approximately well to the true signal. In particular, when the tight frame is an identity matrix or an orthonormal basis, all results obtained in this paper appeared in [13] and [26].

Keywords. Compressed sensing, D -Restricted isometry property, Tight frames, l_q -minimization, Sparse recovery, Coherence.

1 Introduction

Compressed sensing is a new type of sampling theory, that predicts sparse signals can be reconstructed from what was previously believed to be incomplete information [2, 3, 4]. By now, applications of compressed sensing are abundant and range from medical imaging and error correction to radar and remote sensing, see [5, 6] and the references therein.

In compressed sensing, one considers the following model:

$$y = Ax + z, \tag{1.1}$$

where A is a known $m \times n$ measurement matrix (with $m \ll n$) and $z \in \mathbb{R}^n$ is a vector of measurement errors. The goal is to reconstruct the unknown signal x based on y and A . The key idea of compressed sensing relies on that signal is sparse or approximately sparse. A naive approach for

*This work is supported by NSF of China under grant numbers 10771190, 10971189 and Zhejiang Provincial NSF of China under grant number Y6090091.

[†]Corresponding author: Junhong Lin.

E-mail address: songli@zju.edu.cn (S. Li), jhlin5@hotmail.com (J. Lin).

2010 Mathematics Subject Classification. Primary 94A12, 94A15, 94A08, 68P30; Secondary 41A63, 15B52, 42C15.

solving this problem consists in searching for the sparsest vector that is consistent with the linear measurements, which leads to:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_0 \quad \text{subject to} \quad \|A\tilde{x} - y\|_2 \leq \varepsilon, \quad (L_{0,\varepsilon})$$

where $\|x\|_0$ is the numbers of nonzero components of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\|\cdot\|_2$ denotes the standard Euclidean norm and $\varepsilon \geq 0$ is a likely upper bound on the noise level $\|z\|_2$. If $\varepsilon = 0$, it is for the noiseless case. If $\varepsilon > 0$, it is for the noisy case. We call that a vector x is s -sparse if $\|x\|_0 \leq s$. Unfortunately, solving $(L_{0,\varepsilon})$ directly is NP-hard in general and thus is computationally infeasible [7, 8]. One of the practical and tractable alternatives to $(L_{0,\varepsilon})$ proposed in the literature is:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_1 \quad \text{subject to} \quad \|A\tilde{x} - y\|_2 \leq \varepsilon, \quad (L_{1,\varepsilon})$$

which is a convex optimization problem and can be seen as a convex relaxation of $(L_{0,\varepsilon})$. The restricted isometry property (RIP), which first appeared in [9], is one of the most commonly used frameworks for sparse recovery via $(L_{1,\varepsilon})$. For an integer s with $1 \leq s \leq n$, we define the s -restricted isometry constants of a matrix A as the smallest constants satisfying

$$(1 - \delta_s)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_s)\|x\|_2^2$$

for all s -sparse vectors x in \mathbb{R}^n . By computation, one would observe that

$$\delta_s = \max_{T \subset \{1, \dots, n\}, |T| \leq s} \|A_T^* A_T - I\|, \quad (1.2)$$

where $\|\cdot\|$ denotes the spectral norm of a matrix. However, it would be computationally difficult to compute δ_s using (1.2). One of the good news is that many types of random measurement matrices have small restricted isometry constants with very high probability provided that the measurements m is large enough [3, 23, 24]. Since $\delta_{2s} < 1$ is a necessary and sufficient condition to guarantee that any s -sparse vector f is exactly recovered via $(L_{0,0})$ in the noiseless case, many attentions have been focused on δ_{2s} in the literature [10, 11, 12, 13]. Candès [10] showed that under the condition $\delta_{2s} < 0.414$, one can recover a (approximately) sparse signal with a small or zero error using $(L_{1,\varepsilon})$. Later, the sufficient condition on δ_{2s} were improved to $\delta_{2s} < 0.453$ by Lai et al. [11] and $\delta_{2s} < 0.472$ by Cai et al. [12], respectively. Recently, Li and Mo [13] has improved the sufficient condition to $\delta_{2s} < 0.493$ and for some special cases, to $\delta_{2s} < 0.656$. To the best of our knowledge, this is the best known bound on δ_{2s} in the literature. On the other hand, Davies and Gribonval [20] constructed examples which showed that if $\delta_{2s} \geq 1/\sqrt{2}$, exact recovery of certain s -sparse signals can fail in the noiseless case.

For signals which are sparse in the standard coordinate basis or sparse in terms of some other orthonormal basis, the mechanism above holds. However, in practical examples, there are numerous signals of interest which are not sparse in an orthonormal basis. More often than not, sparsity is not expressed in terms of an orthogonal basis but in terms of an overcomplete dictionary [1].

In this paper, we consider recovery of signals that are sparse in terms of a tight frame from undersampled data. Formally, let D be a $n \times d$ matrix whose d columns D_1, \dots, D_d form a tight

frame for \mathbb{R}^n , i.e.

$$f = \sum_k \langle f, D_k \rangle D_k \quad \text{and} \quad \|f\|_2^2 = \sum_k |\langle f, D_k \rangle|^2 \quad \text{for all } f \in \mathbb{R}^n,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product. Our object in this paper is to reconstruct the unknown signal $f \in \mathbb{R}^n$ from a collection of m linear measurements $y = Af + z$ under the assumption that D^*f is sparse or nearly sparse. Such problem has been considered in [22, 25, 14, 1]. The methods in [22, 14, 25] force incoherence on the dictionary D so that the matrix AD conforms to the above standard compressed sensing results. As a result, they may not be suitable for dictionary which are largely correlated. One new alternative way imposing no such properties on the dictionary D for reconstructing the signal f from $y = Af + z$ is to find the solution of l_1 -minimization:

$$\hat{f} = \arg \min_{\tilde{f} \in \mathbb{R}^n} \|D^* \tilde{f}\|_1 \quad \text{subject to} \quad \|A\tilde{f} - y\|_2 \leq \varepsilon, \quad (P_{1,\varepsilon})$$

where again $\varepsilon \geq 0$ is a likely upper bound on the noise level $\|z\|_2$. For discussion of the performance of this method, we would like to introduce the definition of D -RIP of a measurement matrix, which first appeared in [1] and is a natural extension to the standard RIP.

Definition 1.1 (D -RIP). *Let D be a tight frame and Σ_s be the set of all s -sparse vectors in \mathbb{R}^d . A measurement matrix A is said to obey the restricted isometry property adapted to D (abbreviated D -RIP) with constants δ_s if*

$$(1 - \delta_s) \|Dv\|_2^2 \leq \|ADv\|_2^2 \leq (1 + \delta_s) \|Dv\|_2^2$$

holds for all $v \in \Sigma_s$.

For the rest of this paper, D is a $n \times d$ tight frame and δ_s denotes the D -RIP constants with order s of the measurement matrix A without special mentioning. Throughout this paper, denote $x_{[s]}$ to be the vector consisting of the s -largest coefficients of $v \in \mathbb{R}^d$ in magnitude:

$$x_{[s]} = \arg \min_{\|\tilde{x}\|_0 \leq s} \|x - \tilde{x}\|_2.$$

Candès et al. [1] showed that if A satisfies D -RIP with $\delta_{2s} < 0.08$ (in fact, a weaker condition $\delta_{7s} < 0.6$ was stated), then the solution \hat{f} to $(P_{1,\varepsilon})$ satisfies

$$\|\hat{f} - f\|_2 \leq C_0 \frac{\|D^*f - (D^*f)_{[s]}\|_1}{\sqrt{s}} + C_1\varepsilon, \quad (1.3)$$

where the constants C_0 and C_1 may only depend on δ_{2s} . It is easy to see that it's computationally difficult to verify the D -RIP for a given deterministic matrix. But for matrices with Gaussian, sub-gaussian, or Bernoulli entries, the D -RIP condition will be satisfied with overwhelming probability provided that the numbers of measurements m is on the order of $s \log(d/s)$. In fact, for any $m \times n$ matrix A obeying for any fixed $\nu \in \mathbb{R}^n$,

$$\mathbb{P} \left(\left| \|A\nu\|_2^2 - \|\nu\|_2^2 \right| \geq \delta \|\nu\|_2^2 \right) \leq ce^{-\gamma m \delta^2}, \quad \delta \in (0, 1) \quad (1.4)$$

(γ, c are positive numerical constants) will satisfy the D -RIP with overwhelming probability provided that $m \gtrsim s \log(d/s)$ [1]. Therefore, by using D -RIP, the work is independent on the coherence of the dictionary. The result holds even when the coherence of the dictionary D is maximal, meaning two columns are completely correlated. Although Canès et al. in [1] gave the sufficient condition on δ_{2s} to guarantee approximately recovery of a signal via $(P_{1,\varepsilon})$, the bound on δ_{2s} is much weaker comparing to the case for which D is an orthonormal basis. We focus on improving it in this paper.

Our first goal of this paper is to show that the sufficient condition on δ_{2s} above can be improved to $\delta_{2s} < 0.493$. And in some special cases, the sufficient condition can be improved to $\delta_{2s} < 0.656$. These results are given and proved in Section 3. Weakening the D -RIP condition has several benefits. First, it allows more measurement matrices to be used in compressed sensing. Secondly, it give better error estimation in a general problem to recover noisy compressible signal. For example, if $\delta_{2s} = 1/14$, Then by [27, Corollary 3.4], $\delta_{7s} \leq 0.5$. Using the approach in [1] one would get a estimation in (1.3) with $C_0 = 30, C_1 = 62$. While by Theorem 3.4 in Section 3 of this paper, one would get a estimation in (1.3) with $C_0 \simeq 5.06$ and $C_1 \simeq 10.57$. Finally, for the same measurement random matrix A which satisfies (1.4), a standard argument as in [24, 1] shows that it allows recovering a sparse signal with more non-zero transform coefficients. In a nutshell, weakening the D -RIP condition for $(P_{1,\varepsilon})$ is as important as weakening the RIP condition for classical $(L_{1,\varepsilon})$.

Note that the l_0 -norm is the limit as $q \rightarrow 0$ of the l_q -norm in the following sense:

$$\|x\|_0 = \lim_{q \rightarrow 0} \|x\|_q^q = \lim_{q \rightarrow 0} \sum_j |x_j|^q.$$

Thus l_q -norm with $0 < q < 1$ can be used for measuring sparsity. Therefore, one alternative way of finding the solution of $(L_{0,\varepsilon})$ proposed in the literature is to solve:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_q \quad \text{subject to} \quad \|A\tilde{x} - y\|_2 \leq \varepsilon. \quad (L_{q,\varepsilon})$$

This is a non-convex optimization problem since l_q -norm with $0 < q < 1$ is not a norm but a quasi-norm. For any fixed $0 < q < 1$, while checking the global minimal value of $(L_{q,\varepsilon})$ is NP-hard, computing a local minimizer of the problem is polynomial time doable [15]. Therefore, to solve $(L_{q,\varepsilon})$ is still much faster than to solve $(L_{0,\varepsilon})$ at least locally. Reconstruction sparse signals via $(L_{q,\varepsilon})$ with $0 < q < 1$ have been considered in the literature in a series of papers [17, 16, 18, 11, 19, 26] and some of the virtues are highlighted recently. Lai and Liu [26] showed that as long as the classical restricted isometry constant $\delta_{2s} < 1/2$, there exist a value $q_0 \in (0, 1]$ such that for any $q \in (0, q_0)$, each solution of $(L_{q,0})$ for the sparse solution of any underdetermined linear system is the sparsest solution. Thus, it's natural for us to consider the reconstruction of a signal f from $y = Af + z$ by the method of l_q -minimization ($0 < q < 1$):

$$\hat{f} = \arg \min_{\tilde{f} \in \mathbb{R}^n} \|D^* \tilde{f}\|_q \quad \text{subject to} \quad \|A\tilde{f} - y\|_2 \leq \varepsilon. \quad (P_{q,\varepsilon})$$

Our second goal of this paper is to estimate the approximately error between \hat{f} and f when using the l_q -minimization $(P_{q,\varepsilon})$. We show that if the measurement matrix satisfies the D -RIP condition

with $\delta_{2s} < 1/2$, then there exists a value $q_0 = q_0(\delta_{2s}) \in (0, 1]$ such that for any $q \in (0, q_0)$, each solution of the l_q -minimization is approximately well to the true signal f .

This paper is organized as follows. Some lemmas and notations are introduced in Section 2. Section 3 is devoted to discuss recovery of a signal from noisy data via l_1 -minimization. We begin by give some lemmas in this section. And then we discuss approximately recovery of a signal for the general case in Subsection 3.1 while for the special case in Subsection 3.2. Our main results in this section are Theorem 3.4, Theorem 3.8. In Section 4, we discuss recovery of a signal from noisy data via l_q -minimization with $0 < q \leq 1$. Some lemmas and notations are introduced at the beginning. Subsequently, we give the main result Theorem 4.4 and prove it.

2 Lemmas

We will give some lemmas and notations first. We begin by discussing some of the results of recovery of a signal by l_0 -minimization:

$$\hat{f} = \arg \min_{\tilde{f} \in \mathbb{R}^n} \|D^* \tilde{f}\|_0, \quad \text{subject to} \quad A\tilde{f} = y. \quad (P_0)$$

Lemma 2.1. *If $\delta_{2s} < 1$, then every signal f such that $\|D^* f\|_0 \leq s$ can be exactly recovered by solving (P_0) .*

Proof. Since \hat{f} is the solution of (P_0) , we have $\|D^* \hat{f}\|_0 \leq \|D^* f\|_0 \leq s$. Then $h = \hat{f} - f$ satisfies $\|D^* h\|_0 \leq 2s$. Notice that $0 = Ah = ADD^* h$. Thus according to the definition of D -RIP, we have

$$(1 - \delta_{2s}) \|DD^* h\|_2^2 \leq \|ADD^* h\|_2^2 = 0.$$

Combining with $\delta_{2s} < 1$, we get $0 = DD^* h = h$. The proof is finished. \square

Lemma 2.2. *For all $u, v \in \Sigma_s$, we have*

$$\langle ADu, ADv \rangle \leq \delta_{2s} \|Du\|_2 \|Dv\|_2 + \langle Du, Dv \rangle.$$

Proof. For $u, v \in \Sigma_s$, assume that $\|Du\|_2 = \|Dv\|_2 = 1$. By the definition of D -RIP, we have

$$\begin{aligned} \langle ADu, ADv \rangle &= \frac{1}{4} \{ \|ADu + ADv\|_2^2 - \|ADu - ADv\|_2^2 \} \\ &\leq \frac{1}{4} \{ (1 + \delta_{2s}) \|Du + Dv\|_2^2 - (1 - \delta_{2s}) \|Du - Dv\|_2^2 \} = \delta_{2s} + \langle Du, Dv \rangle. \end{aligned}$$

Thus by a simple modification, we conclude the proof. \square

Remark 2.3. *Notice that when D is an identity matrix and for any s -sparse vectors u, v with disjoint supports, one have*

$$\langle Au, Av \rangle \leq \delta_{2s} \|u\|_2 \|v\|_2,$$

which is the same as [10, Lemma 2.1].

For $T \subset \{1, \dots, d\}$, denote by D_T the matrix D restricted to the columns indexed by T , and write D_T^* to mean $(D_T)^*$, T^c to mean the complement of T in $\{1, \dots, d\}$. Given a vector $h \in \mathbb{R}^n$, we write $D^*h = (x_1, \dots, x_s, \dots, x_{2s}, \dots, x_d)$. Let $D^*h = D_{T_0}^*h + D_{T_1}^*h + \dots + D_{T_l}^*h$, where $D_{T_0}^*h = (x_1, x_2, \dots, x_s, 0, 0, \dots, 0)$, $D_{T_1}^*h = (0, 0, \dots, 0, x_{s+1}, x_{s+2}, \dots, x_{2s}, 0, 0, \dots, 0)$, $D_{T_2}^*h = (0, 0, \dots, 0, x_{2s+1}, x_{2s+2}, \dots, x_{3s}, 0, 0, \dots, 0)$, \dots and $D_{T_l}^*h = (0, 0, \dots, 0, x_{ls+1}, x_{ls+2}, \dots, x_d)$. Denote $T_{01} = T_0 \cup T_1$ and $T = T_0$. For simplicity, we assume that $\delta_{2s} < 1$.

Lemma 2.4. *We have*

$$\left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 \leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2.$$

Proof. By the definition of δ_{2s} and Lemma 2.2, we have

$$\begin{aligned} & \left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 = \sum_{i,j=2}^l \langle ADD_{T_i}^* h, ADD_{T_j}^* h \rangle \\ &= \sum_{j=2}^l \|ADD_{T_j}^* h\|_2^2 + 2 \sum_{2 \leq i < j \leq l} \langle ADD_{T_i}^* h, ADD_{T_j}^* h \rangle \\ &\leq (1 + \delta_{2s}) \sum_{j=2}^l \|DD_{T_j}^* h\|_2^2 + 2 \sum_{2 \leq i < j \leq l} \langle DD_{T_i}^* h, DD_{T_j}^* h \rangle + 2\delta_{2s} \sum_{2 \leq i < j \leq l} \|DD_{T_i}^* h\|_2 \|DD_{T_j}^* h\|_2 \\ &= \left\| \sum_{j=2}^l DD_{T_j}^* h \right\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|DD_{T_j}^* h\|_2 \right)^2. \end{aligned}$$

Notice that by $\|h\|_2^2 = \sum_{j=0}^l \|D_{T_j}^* h\|_2^2$,

$$\begin{aligned} & \left\| \sum_{j=2}^l DD_{T_j}^* h \right\|_2^2 = \|h - DD_{T_{01}}^* h\|_2^2 = \|h\|_2^2 - 2\langle h, DD_{T_{01}}^* h \rangle + \|DD_{T_{01}}^* h\|_2^2 \\ &= \|h\|_2^2 - 2\|D_{T_{01}}^* h\|_2^2 + \|DD_{T_{01}}^* h\|_2^2 = \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 - \|D_{T_{01}}^* h\|_2^2 + \|DD_{T_{01}}^* h\|_2^2. \end{aligned} \quad (2.1)$$

Therefore, we get

$$\begin{aligned} \left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 &\leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 - \|D_{T_{01}}^* h\|_2^2 + \|DD_{T_{01}}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 \\ &\leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2, \end{aligned} \quad (2.2)$$

where we have used $\|DD_{T_j}^* h\|_2 \leq \|D_{T_j}^* h\|_2, j \in \{0, 1, \dots, l\}$. □

Lemma 2.5. *We have*

$$\left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 - \left\| ADD_{T_{01}}^* h \right\|_2^2 \leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 - (1 - \delta_{2s}) \|D_{T_{01}}^* h\|_2^2.$$

Proof. By the definition of δ_{2s} and (2.2), we have

$$\begin{aligned} & \left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 - \left\| ADD_{T_{01}}^* h \right\|_2^2 \\ & \leq \left\| \sum_{j=2}^l ADD_{T_j}^* h \right\|_2^2 - (1 - \delta_{2s}) \|DD_{T_{01}}^* h\|_2^2 \\ & \leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 + \delta_{2s} \|DD_{T_{01}}^* h\|_2^2 - \|D_{T_{01}}^* h\|_2^2 \\ & \leq \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 - (1 - \delta_{2s}) \|D_{T_{01}}^* h\|_2^2. \end{aligned}$$

□

3 Recovery via l_1 -minimization

In this section, we are concerned with the reconstruction of a signal f from $y = Af + z$ by the method of l_1 -minimization: $(P_{1,\varepsilon})$.

Let $h = \hat{f} - f$, where \hat{f} is the solution of $(P_{1,\varepsilon})$ and f is the original signal. We use the same assumptions as in Section 2. Furthermore, rearranging the indices if necessary, we assume that the first s coordinates of D^*f are the largest in magnitude and $|x_{s+1}| \geq |x_{s+2}| \geq \dots \geq |x_d|$. For the rest of this section, we will always assume that $\|D_{T_1}^* h\|_1 = \omega \sum_{j=1}^l \|D_{T_j}^* h\|_1$ for some nonnegative number $\omega \in [0, 1]$. Then we have $\sum_{j=2}^l \|D_{T_j}^* h\|_1 = (1 - \omega) \sum_{j=1}^l \|D_{T_j}^* h\|_1$.

Lemma 3.1. *We have*

$$\sum_{j=2}^l \|D_{T_j}^* h\|_2^2 \leq \frac{\omega(1 - \omega)}{s} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_1 \right)^2.$$

Proof. By the simple inequality

$$\sum_{j=1}^d |x_j|^2 \leq \max_{1 \leq j \leq d} |x_j| \sum_{j=1}^d |x_j|,$$

we have

$$\sum_{j=2}^l \|D_{T_j}^* h\|_2^2 \leq |x_{2s+1}| \sum_{j=2}^l \|D_{T_j}^* h\|_1 \leq \frac{\omega(1-\omega)}{s} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_1 \right)^2.$$

□

Lemma 3.2. *We have*

$$\sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 \leq \frac{\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2}{s} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_1 \right)^2.$$

Proof. By [21, Proposition 1], we have

$$s^{1/2} \|D_{T_j}^* h\|_2 \leq \|D_{T_j}^* h\|_1 + s(|x_{js+1}| - |x_{js+s}|)/4, \quad j = 2, \dots, l.$$

Therefore, we have

$$s^{1/2} \sum_{j=2}^l \|D_{T_j}^* h\|_2 \leq \sum_{j=2}^l \|D_{T_j}^* h\|_1 + s|x_{2s+1}|/4 \leq \sum_{j=2}^l \|D_{T_j}^* h\|_1 + \|D_{T_1}^* h\|_1/4 = (1-3\omega/4) \sum_{j=1}^l \|D_{T_j}^* h\|_1.$$

Combining the above inequality with Lemma 3.1, one can finish the proof. □

Since \hat{f} is a minimizer of $(P_{1,\varepsilon})$, one gets that

$$\|D^* f\|_1 \geq \|D^* \hat{f}\|_1.$$

That is

$$\|D_T^* f\|_1 + \|D_{T^c}^* f\|_1 \geq \|D_T^* \hat{f}\|_1 + \|D_{T^c}^* \hat{f}\|_1.$$

Thus

$$\|D_T^* f\|_1 + \|D_{T^c}^* f\|_1 \geq \|D_T^* f\|_1 - \|D_T^* h\|_1 + \|D_{T^c}^* h\|_1 - \|D_{T^c}^* f\|_1.$$

This implies

$$\sum_{j=1}^l \|D_{T_j}^* h\|_1 \leq 2\|D_{T^c}^* f\|_1 + \|D_T^* h\|_1. \quad (3.1)$$

According to the feasibility of \hat{f} , Ah must be small:

$$\|Ah\|_2 = \|Af - A\hat{f}\|_2 \leq \|Af - y\|_2 + \|A\hat{f} - y\|_2 \leq 2\varepsilon. \quad (3.2)$$

3.1 General signal recovery

For $\delta_{2s} < 2/3$, denote that

$$\rho_s = \sqrt{\frac{4(1 + 5\delta_{2s} - 4\delta_{2s}^2)}{(1 - \delta_{2s})(32 - 25\delta_{2s})}}. \quad (3.3)$$

By a easy computation, one can show that if $\delta_{2s} < (77 - \sqrt{1337})/82 \approx 0.4931$, then $\rho_s < 1$.

Lemma 3.3. *If $\delta_{2s} < 0.4931$, then*

$$\sum_{j=1}^l \|D_{T_j}^* h\|_1 \leq \frac{2}{1 - \rho_s} \|D^* f - (D^* f)_{[s]}\|_1 + \frac{2\sqrt{2}}{(1 - \rho_s)\sqrt{1 - \delta_{2s}}} \sqrt{s}\varepsilon. \quad (3.4)$$

Proof. By $Ah = \sum_{j=0}^l ADD_{T_j}^* h$ and (3.2), we have

$$\begin{aligned} 0 &= \|Ah - \sum_{j=2}^l ADD_{T_j}^* h\|_2^2 - \|ADD_{T_{01}}^* h\|_2^2 \\ &\leq (2\varepsilon + \|\sum_{j=2}^l ADD_{T_j}^* h\|_2)^2 - \|ADD_{T_{01}}^* h\|_2^2 \\ &= 4\varepsilon^2 + 4\varepsilon \|\sum_{j=2}^l ADD_{T_j}^* h\|_2 + \|\sum_{j=2}^l ADD_{T_j}^* h\|_2^2 - \|ADD_{T_{01}}^* h\|_2^2 \end{aligned}$$

Applying lemmas 2.4 and 2.5 to the above inequality yields

$$\|D_{T_{01}}^* h\|_2^2 \leq \frac{(2\varepsilon + \mathcal{N})^2}{1 - \delta_{2s}}, \quad (3.5)$$

where

$$\mathcal{N} = \sqrt{\sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2}.$$

Using

$$\|D_{T_{01}}^* h\|_2^2 = \|D_{T_0}^* h\|_2^2 + \|D_{T_1}^* h\|_2^2 \geq \frac{\|D_{T_0}^* h\|_1^2 + \|D_{T_1}^* h\|_1^2}{s} \quad (3.6)$$

to (3.5), one can get

$$\begin{aligned} \|D_{T_0}^* h\|_1^2 &\leq \frac{s(2\varepsilon + \mathcal{N})^2 - (1 - \delta_{2s})\|D_{T_1}^* h\|_1^2}{1 - \delta_{2s}} \\ &\leq \frac{(2\sqrt{2s}\varepsilon)^2 + 2 \cdot 2\sqrt{2s}\varepsilon \cdot \sqrt{s/2}\mathcal{N} + s\mathcal{N}^2 - (1 - \delta_{2s})\|D_{T_1}^* h\|_1^2}{1 - \delta_{2s}}. \end{aligned} \quad (3.7)$$

Notice that by Lemma 3.2, we have

$$\begin{aligned}\sqrt{\frac{s}{2}}\mathcal{N} &\leq \sqrt{\frac{\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2}{2}} \sum_{j=1}^l \|D_{T_j}^* h\|_1 \\ &\leq \sqrt{\frac{4(1+5\delta_{2s}-4\delta_{2s}^2)}{(32-25\delta_{2s})}} \sum_{j=1}^l \|D_{T_j}^* h\|_1\end{aligned}$$

and

$$\begin{aligned}s\mathcal{N}^2 - (1-\delta_{2s})\|D_{T_1}^* h\|_1^2 &\leq [\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2 - (1-\delta_{2s})\omega^2] \left(\sum_{j=1}^l \|D_{T_j}^* h\|_1 \right)^2 \\ &\leq \frac{4(1+5\delta_{2s}-4\delta_{2s}^2)}{(32-25\delta_{2s})} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_1 \right)^2,\end{aligned}$$

where we have used the fact for all $\delta_{2s} \in [0, 2/3]$,

$$\begin{aligned}\max_{\omega \in [0,1]} \frac{\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2}{2} &\leq \frac{\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2}{2} \Big|_{\omega = \frac{4(2-3\delta_{2s})}{16-9\delta_{2s}}} \\ &= \frac{2(1+\delta_{2s})}{16-9\delta_{2s}} \leq \frac{4(1+5\delta_{2s}-4\delta_{2s}^2)}{(32-25\delta_{2s})}\end{aligned}$$

and

$$\max_{\omega \in [0,1]} \omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2 - (1-\delta_{2s})\omega^2 \leq \frac{4(1+5\delta_{2s}-4\delta_{2s}^2)}{(32-25\delta_{2s})}.$$

Thus, it follows from the above two inequalities and (3.7) that

$$\|D_T^* h\|_1 \leq \frac{2\sqrt{2s}\varepsilon}{\sqrt{1-\delta_{2s}}} + \rho_s \sum_{j=1}^l \|D_{T_j}^* h\|_1.$$

Combining with (3.1), one can finish the proof. \square

The main result of this subsection is the following theorem.

Theorem 3.4. *If $\delta_{2s} < 0.4931$, then*

$$\|\hat{f} - f\|_2 \leq C_0 \frac{\|D^* f - (D^* f)_{[s]}\|_1}{\sqrt{s}} + C_1 \varepsilon, \quad (3.8)$$

where

$$C_0 = \frac{4}{1-\rho_s} \sqrt{\frac{2(2-\delta_{2s})}{(1-\delta_{2s})(32-25\delta_{2s})}}, \quad C_1 = \frac{2}{\sqrt{1-\delta_{2s}}} \left(1 + \frac{C_0}{\sqrt{2}} \right)$$

and

$$\rho_s = \sqrt{\frac{4(1+5\delta_{2s}-4\delta_{2s}^2)}{(1-\delta_{2s})(32-25\delta_{2s})}}.$$

Proof. By (3.5), we have

$$\begin{aligned} \|h\|_2^2 &= \|D^*h\|_2^2 = \|D_{T_{01}}^*h\|_2^2 + \sum_{j=2}^l \|D_{T_j}^*h\|_2^2 \\ &\leq \frac{1}{1-\delta_{2s}} \left(2\varepsilon + \sqrt{\sum_{j=2}^l \|D_{T_j}^*h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^*h\|_2 \right)^2} \right)^2 + \sum_{j=2}^l \|D_{T_j}^*h\|_2^2. \end{aligned}$$

Hence, we have

$$\|h\|_2 \leq \frac{2\varepsilon}{\sqrt{1-\delta_{2s}}} + \sqrt{\frac{1}{(1-\delta_{2s})} \left[\sum_{j=2}^l \|D_{T_j}^*h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^*h\|_2 \right)^2 \right]} + \sum_{j=2}^l \|D_{T_j}^*h\|_2. \quad (3.9)$$

Then by lemmas 3.1 and 3.2, we have

$$\|h\|_2 \leq \frac{2\varepsilon}{\sqrt{1-\delta_{2s}}} + \frac{\sqrt{(2-\delta_{2s})\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2}}{\sqrt{s}\sqrt{(1-\delta_{2s})}} \sum_{j=1}^l \|D_{T_j}^*h\|_1.$$

A direct calculation shows that

$$\max_{0 < \omega \leq 1} (2-\delta_{2s})\omega(1-\omega) + \delta_{2s}(1-3\omega/4)^2 \leq \frac{8(2-\delta_{2s})}{32-25\delta_{2s}}.$$

It follows from the above two inequalities that

$$\|h\|_2 \leq \frac{2\varepsilon}{\sqrt{1-\delta_{2s}}} + \frac{2}{\sqrt{s}} \sqrt{\frac{2(2-\delta_{2s})}{(1-\delta_{2s})(32-25\delta_{2s})}} \sum_{j=1}^l \|D_{T_j}^*h\|_1.$$

Therefore, with the above inequality and Lemma 3.3 we prove the result. \square

3.2 Special case: $n \leq 4s$

For $\delta_{2s} \in [0, 1)$, denote that

$$\rho_s = \sqrt{\frac{(1+\delta_{2s})^2}{8(1-\delta_{2s})}}. \quad (3.10)$$

By a easy computation, one can show that if $\delta_{2s} < 4\sqrt{2} - 5 \approx 0.656$, $\rho_s < 1$.

We have $l \leq 3$ by $n \leq 4s$. For simplicity, we assume that $l = 3$. Instead of lemmas 2.4 and 2.5, we have the following results.

Lemma 3.5. *We have*

$$\|ADD_{T_{23}}^*h\|_2^2 \leq (1+\delta_{2s})\|DD_{T_{23}}^*h\|_2^2 \leq (1+\delta_{2s})\|D_{T_{23}}^*h\|_2^2.$$

Proof. The proof is straightforward. \square

Lemma 3.6. *We have*

$$\|ADD_{T_{23}}^* h\|_2^2 - \|ADD_{T_{01}}^* h\|_2^2 \leq (1 + \delta_{2s})\|D_{T_{23}}^* h\|_2^2 - (1 - \delta_{2s})\|D_{T_{01}}^* h\|_2^2.$$

Proof. By the definition of δ_{2s} , we have

$$\|ADD_{T_{23}}^* h\|_2^2 - \|ADD_{T_{01}}^* h\|_2^2 \leq (1 + \delta_{2s})\|DD_{T_{23}}^* h\|_2^2 - (1 - \delta_{2s})\|DD_{T_{01}}^* h\|_2^2. \quad (3.11)$$

Applying (2.1) to the above yields

$$\begin{aligned} \|ADD_{T_{23}}^* h\|_2^2 - \|ADD_{T_{01}}^* h\|_2^2 &\leq (1 + \delta_{2s})\|D_{T_{23}}^* h\|_2^2 - (1 + \delta_{2s})\|D_{T_{01}}^* h\|_2^2 + 2\delta_{2s}\|DD_{T_{01}}^* h\|_2^2 \\ &\leq (1 + \delta_{2s})\|D_{T_{23}}^* h\|_2^2 - (1 - \delta_{2s})\|D_{T_{01}}^* h\|_2^2. \end{aligned}$$

\square

Similar as Lemma 3.3, we have the following result.

Lemma 3.7. *If $n \leq 4s$ and $\delta_{2s} < 0.656$, then*

$$\sum_{j=1}^3 \|D_{T_j}^* h\|_1 \leq \frac{2}{1 - \rho_s} \|D^* f - (D^* f)_{[s]}\|_1 + \frac{2\sqrt{2}}{(1 - \rho_s)\sqrt{1 - \delta_{2s}}} \sqrt{s\varepsilon}. \quad (3.12)$$

Proof. Applying lemmas 3.5 and 3.6, and a similar approach as that for (3.7) yields

$$\|D_{T_0}^* h\|_1^2 \leq \frac{(2\sqrt{2s\varepsilon})^2 + 2 \cdot 2\sqrt{2s\varepsilon} \cdot \sqrt{s/2}\mathcal{N} + s\mathcal{N}^2 - (1 - \delta_{2s})\|D_{T_1}^* h\|_1^2}{1 - \delta_{2s}}, \quad (3.13)$$

where

$$\mathcal{N} = \sqrt{(1 + \delta_{2s})} \|D_{T_{23}}^* h\|_2.$$

Notice that by applying Lemma 3.1, we have

$$\sqrt{\frac{s}{2}}\mathcal{N} \leq \sqrt{\frac{\omega(1 - \omega)(1 + \delta_{2s})}{2}} \sum_{j=1}^3 \|D_{T_j}^* h\|_1 \leq \sqrt{\frac{1 + \delta_{2s}}{8}} \sum_{j=1}^3 \|D_{T_j}^* h\|_1 \leq \sqrt{\frac{(1 + \delta_{2s})^2}{8}} \sum_{j=1}^3 \|D_{T_j}^* h\|_1$$

and

$$\begin{aligned} s\mathcal{N}^2 - (1 - \delta_{2s})\|D_{T_1}^* h\|_1^2 &\leq [\omega(1 - \omega)(1 + \delta_{2s}) - (1 - \delta_{2s})\omega^2] \left(\sum_{j=1}^3 \|D_{T_j}^* h\|_1 \right)^2 \\ &\leq \frac{(1 + \delta_{2s})^2}{8} \left(\sum_{j=1}^3 \|D_{T_j}^* h\|_1 \right)^2. \end{aligned}$$

Thus, it follows from the above two inequalities and (3.13) that

$$\|D_T^* h\|_1 \leq \frac{2\sqrt{2s\varepsilon}}{\sqrt{1 - \delta_{2s}}} + \rho_s \sum_{j=1}^3 \|D_{T_j}^* h\|_1.$$

Combining with (3.1), one can finish the proof. \square

The main result of this subsection is the following theorem.

Theorem 3.8. *If $n \leq 4s$ and $\delta_{2s} < 0.656$, then*

$$\|\hat{f} - f\|_2 \leq C_0 \frac{\|D^*f - (D^*f)_{[s]}\|_1}{\sqrt{s}} + C_1\varepsilon, \quad (3.14)$$

where

$$C_0 = \frac{\sqrt{2}}{(1 - \rho_s)\sqrt{(1 - \delta_{2s})}}, \quad C_1 = \frac{2}{\sqrt{1 - \delta_{2s}}} \left(1 + \frac{C_0}{\sqrt{2}}\right) \quad \text{and} \quad \rho_s = \sqrt{\frac{(1 + \delta_{2s})^2}{8(1 - \delta_{2s})}}.$$

Proof. By a similar approach as that for (3.9), we have

$$\|h\|_2 \leq \frac{2\varepsilon}{\sqrt{1 - \delta_{2s}}} + \sqrt{\frac{1}{(1 - \delta_{2s})}(1 + \delta_{2s})\|D_{T_{23}}^*h\|_2^2 + \|D_{T_{23}}^*h\|_2^2}.$$

Then by lemmas 3.1, we have

$$\|h\|_2 \leq \frac{2\varepsilon}{\sqrt{1 - \delta_{2s}}} + \frac{\sqrt{2\omega(1 - \omega)}}{\sqrt{s}\sqrt{(1 - \delta_{2s})}} \sum_{j=1}^3 \|D_{T_j}^*h\|_1 \leq \frac{2\varepsilon}{\sqrt{1 - \delta_{2s}}} + \frac{1}{\sqrt{s}\sqrt{2(1 - \delta_{2s})}} \sum_{j=1}^3 \|D_{T_j}^*h\|_1.$$

Therefore, with the above inequality and Lemma 3.7 we prove the result. \square

Remark 3.9. *When D is an identity matrix, that is for the classical RIP and l_1 -minimization ($L_{1,\varepsilon}$), theorems 3.4 and 3.8 were proved in [13].*

4 Recovery via l_q -minimization with $0 < q < 1$

In this section, we will discuss recovery of a signal by l_q -minimization ($P_{q,\varepsilon}$) with $0 < q < 1$. For $q \in (0, 1]$, let $h = \hat{f} - f$, where \hat{f} is the solution of ($P_{q,\varepsilon}$) and f is the original signal. We follow the same assumptions as in Section 2. Moreover, rearranging the indices if necessary, we assume that the first s coordinates of D^*f are the largest in magnitude and $|x_{s+1}| \geq |x_{s+2}| \geq \dots \geq |x_d|$. For the rest of this section, for $q \in (0, 1]$, we will always assume that $\|D_{T_1}^*h\|_q^q = \omega \sum_{j=1}^l \|D_{T_j}^*h\|_q^q$ for some nonnegative number $\omega = \omega(q) \in [0, 1]$. Then we have $\sum_{j=2}^l \|D_{T_j}^*h\|_q^q = (1 - \omega) \sum_{j=1}^l \|D_{T_j}^*h\|_q^q$.

Lemma 4.1. *For $q \in (0, 1]$, we have*

$$\sum_{j=2}^l \|D_{T_j}^*h\|_2^2 \leq \frac{(1 - \omega)\omega^{(2-q)/q}}{s^{(2-q)/q}} \left(\sum_{j=1}^l \|D_{T_j}^*h\|_q^q \right)^{2/q}.$$

Proof. By the simple inequality

$$\sum_{j=1}^d |x_j|^2 \leq \max_{1 \leq j \leq d} |x_j|^{2-q} \sum_{j=1}^d |x_j|^q,$$

we have

$$\begin{aligned} \sum_{j=2}^l \|D_{T_j}^* h\|_2^2 &\leq |x_{2s+1}|^{2-q} \sum_{j=2}^l \|D_{T_j}^* h\|_q^q \leq \left(\frac{\|D_{T_1}^* h\|_q^q}{s} \right)^{(2-q)/q} \sum_{j=2}^l \|D_{T_j}^* h\|_q^q \\ &= \frac{(1-\omega)\omega^{(2-q)/q}}{s^{(2-q)/q}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{2/q}. \end{aligned}$$

□

Lemma 4.2. For $q \in (0, 1]$, we have

$$\sum_{j=2}^l \|D_{T_j}^* h\|_2^2 + \delta_{2s} \left(\sum_{j=2}^l \|D_{T_j}^* h\|_2 \right)^2 \leq \frac{(1-\omega)\omega^{(2-q)/q} + \delta_{2s}}{s^{2/q-1}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{2/q}.$$

Proof. By [26, Lemma 2], we have

$$s^{1/q-1/2} \|D_{T_j}^* h\|_2 \leq \|D_{T_j}^* h\|_q + s^{1/q} (|x_{js+1}| - |x_{js+s}|), \quad j = 2, \dots, l.$$

Therefore, we have

$$s^{1/q-1/2} \sum_{j=2}^l \|D_{T_j}^* h\|_2 \leq \sum_{j=2}^l \|D_{T_j}^* h\|_q + s^{1/q} |x_{2s+1}| \leq \sum_{j=2}^l \|D_{T_j}^* h\|_q + \|D_{T_1}^* h\|_q \leq \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{1/q},$$

where for the last inequality we have used $\|n\|_1 \leq \|n\|_q$ for $n \in \mathbb{R}^l$. Combining with Lemma 4.1, one can conclude the result. □

Analogous to (3.1) and (3.2), one can prove that

$$\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \leq 2 \|D_{T^c}^* f\|_q^q + \|D_T^* h\|_q^q \quad (4.1)$$

and

$$\|Ah\|_2 \leq 2\varepsilon. \quad (4.2)$$

Denote that

$$\rho_s(q) = \sqrt{\frac{\delta_{2s}}{1-\delta_{2s}} + \frac{q}{2^{2/q}(1-\delta_{2s})} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1}}.$$

For $\delta_{2s} < \frac{1}{2}$, one can prove that there exists a value $q_0 = q_0(\delta_{2s}) \in (0, 1]$ such that for all $q \in (0, q_0)$, $\rho_s(q) < 1$. Indeed, by a easy calculation, $\rho_s(q) < 1$ is equivalent to

$$\delta_{2s} + \frac{q}{2^{2/q+1}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1} < \frac{1}{2}.$$

Since the second term on the left hand side goes to zero as $q \rightarrow 0_+$ as $\delta_{2s} < 1, q \leq 1$ and

$$\frac{1}{2^{2/q+1}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1} \leq \left(\frac{2-q}{2} \right)^{2/q} \approx \frac{1}{e},$$

one can finish the conclusion.

Lemma 4.3. *If $\delta_{2s} < 1/2$ and $q \in (0, q_0)$, then*

$$\left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}} \leq \frac{2^{\frac{2}{q}-1}}{(1-\rho_s^q(q))^{1/q}} \|D^* f - (D^* f)_{[s]}\|_q + \frac{2^{\frac{2}{q}-\frac{1}{2}} s^{\frac{1}{q}-\frac{1}{2}} \varepsilon}{(1-\rho_s^q(q))^{1/q} \sqrt{(1-\delta_{2s})}}.$$

Proof. Notice that we also have (3.5). Using

$$\|D_{T_{01}}^* h\|_2^2 = \|D_{T_0}^* h\|_2^2 + \|D_{T_1}^* h\|_2^2 \geq \frac{\|D_{T_0}^* h\|_q^2 + \|D_{T_1}^* h\|_q^2}{s^{(2-q)/q}} \quad (4.3)$$

to (3.5), one can get

$$\begin{aligned} \|D_{T_0}^* h\|_q^2 &\leq \frac{s^{(2-q)/q} (2\varepsilon + \mathcal{N})^2 - (1-\delta_{2s}) \|D_{T_1}^* h\|_q^2}{1-\delta_{2s}} \\ &\leq \frac{(2(2s)^{\frac{1}{q}-\frac{1}{2}} \varepsilon)^2 + 2 \cdot 2(2s)^{\frac{1}{q}-\frac{1}{2}} \varepsilon \cdot \left(\frac{s}{2}\right)^{\frac{1}{q}-\frac{1}{2}} \mathcal{N} + s^{\frac{2}{q}-1} \mathcal{N}^2 - (1-\delta_{2s}) \|D_{T_1}^* h\|_q^2}{1-\delta_{2s}}. \end{aligned} \quad (4.4)$$

Notice that by Lemma 4.2, we have

$$\begin{aligned} \left(\frac{s}{2}\right)^{\frac{1}{q}-\frac{1}{2}} \mathcal{N} &\leq \sqrt{\frac{(1-\omega)\omega^{(2-q)/q} + \delta_{2s}}{2^{(2-q)/q}}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}} \\ &\leq \sqrt{\delta_{2s} + \frac{q}{2^{2/q}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}} \end{aligned}$$

and

$$\begin{aligned} s^{\frac{2}{q}-1} \mathcal{N}^2 - (1-\delta_{2s}) \|D_{T_1}^* h\|_q^2 &\leq [\omega^{(2-q)/q} (1-\omega) + \delta_{2s} - (1-\delta_{2s}) \omega^{2/q}] \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{2}{q}} \\ &\leq \left(\delta_{2s} + \frac{q}{2^{2/q}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1} \right) \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{2}{q}}, \end{aligned}$$

where we have used the fact for all $\delta_{2s} \in [0, 1)$,

$$\begin{aligned} \max_{\omega \in [0, 1]} \frac{(1-\omega)\omega^{(2-q)/q} + \delta_{2s}}{2^{(2-q)/q}} &\leq \frac{(1-\omega)\omega^{(2-q)/q} + \delta_{2s}}{2^{(2-q)/q}} \Big|_{\omega=\frac{2-q}{2}} = \frac{\delta_{2s} + \frac{q}{2^{2/q}} (2-q)^{\frac{2}{q}-1}}{2^{(2-q)/q}} \\ &\leq \delta_{2s} + \frac{q}{2^{2/q}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1} \end{aligned}$$

and

$$\begin{aligned} & \max_{\omega \in [0,1]} \omega^{(2-q)/q} (1-\omega) + \delta_{2s} - (1-\delta_{2s})\omega^{2/q} \\ & \leq \omega^{(2-q)/q} (1-\omega) + \delta_{2s} - (1-\delta_{2s})\omega^{2/q} \Big|_{\omega=\frac{2-q}{2(2-\delta_{2s})}} = \delta_{2s} + \frac{q}{2^{2/q}} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1}. \end{aligned}$$

Thus, it follows from the above two inequalities and (4.4) that

$$\|D_T^* h\|_q \leq \frac{2(2s)^{\frac{1}{q}-\frac{1}{2}}\varepsilon}{\sqrt{1-\delta_{2s}}} + \rho_s(q) \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}}.$$

Therefore, we get

$$\|D_T^* h\|_q^q \leq \frac{2^q(2s)^{1-\frac{q}{2}}\varepsilon^q}{(1-\delta_{2s})^{\frac{q}{2}}} + \rho_s^q(q) \sum_{j=1}^l \|D_{T_j}^* h\|_q^q,$$

where we have used $(|a| + |b|)^q \leq |a|^q + |b|^q$. Plug the above inequality to (4.1) and by a easy computation, one can show that

$$\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \leq \frac{2}{1-\rho_s^q(q)} \|D_{T^c}^* f\|_q^q + \frac{2^q(2s)^{1-\frac{q}{2}}\varepsilon^q}{(1-\delta_{2s})^{\frac{q}{2}}(1-\rho_s^q(q))}.$$

Using the basic inequality $\|v\|_q \leq 2^{\frac{1}{q}-1} \|v\|_1$ for $v \in \mathbb{R}^2$ to the above inequality, one can finish the proof. \square

The main result of this subsection is the following theorem.

Theorem 4.4. *If $\delta_{2s} < \frac{1}{2}$, then there exists a value $q_0 = q_0(\delta_{2s}) \in (0, 1]$ such that for any $q \in (0, q_0)$,*

$$\rho_s(q) = \sqrt{\frac{\delta_{2s}}{1-\delta_{2s}} + \frac{q}{2^{2/q}(1-\delta_{2s})} \left(\frac{2-q}{2-\delta_{2s}} \right)^{\frac{2}{q}-1}} < 1.$$

Furthermore,

$$\|\hat{f} - f\|_2 \leq C_0 \frac{\|D^* f - (D^* f)_{[s]}\|_q}{s^{\frac{1}{q}-\frac{1}{2}}} + C_1 \varepsilon, \quad (4.5)$$

where

$$C_0 = \frac{2^{1/q-1}}{(1-\rho_s^q(q))^{1/q}} \sqrt{\frac{(2-\delta_{2s})(2-q)^{(2-q)/q}q + 2^{2/q}\delta_{2s}}{1-\delta_{2s}}} \quad \text{and} \quad C_1 = \frac{2}{\sqrt{1-\delta_{2s}}} \left(1 + \frac{C_0}{\sqrt{2}} \right).$$

Proof. Notice that we also have (3.9). Using lemmas 4.1 and 4.2 to (3.9), we get

$$\begin{aligned} \|h\|_2 & \leq \frac{2\varepsilon}{\sqrt{1-\delta_{2s}}} + \sqrt{\frac{(2-\delta_{2s})\omega^{(2-q)/q}(1-\omega) + \delta_{2s}}{(1-\delta_{2s})s^{(2-q)/q}}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}} \\ & \leq \frac{2\varepsilon}{\sqrt{1-\delta_{2s}}} + \sqrt{\frac{(2-\delta_{2s})(2-q)^{(2-q)/q}q + 2^{2/q}\delta_{2s}}{(1-\delta_{2s})2^{2/q}s^{(2-q)/q}}} \left(\sum_{j=1}^l \|D_{T_j}^* h\|_q^q \right)^{\frac{1}{q}}, \end{aligned}$$

where we have used

$$\max_{\omega \in [0,1]} (2 - \delta_{2s})(1 - \omega)\omega^{(2-q)/q} \leq (2 - \delta_{2s})(1 - \omega)\omega^{(2-q)/q} \Big|_{\omega = \frac{2-q}{2}}.$$

Now the proof can be finished by directly applying Lemma 4.3. \square

Remark 4.5. When D is an identity matrix, that is for the classical RIP and l_q -minimization ($L_{q,\varepsilon}$), Theorem 4.4 was shown in [26].

References

- [1] E. J. Candès, Y. C. Eldar and D. Needell, *Compressed Sensing with Coherent and Redundant Dictionaries*, Appl. Comput. Harmon. Anal. (2010), doi:10.1016/j.acha.2010.10.002.
- [2] E. J. Candès, J. Romberg and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, **52** (2006), 489-509.
- [3] E. J. Candès, J. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math., **59** (2006), 1207-1223.
- [4] D. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, **52** (2006), 1289-1306.
- [5] A. M. Zoubir and D. R. Iskander, *Bootstrap methods in Signal Processing*, IEEE Signal Proc. Mag., **24** (2007).
- [6] R. Baraniuk, E. J. Candès, R. Nowak and M. Vetterli, *Sensing, sampling, and compression*, IEEE Signal Proc. Mag., **25**, (2008).
- [7] S. G. Mallat and Z. Zhang, *Matching pursuits with time-frequency dictionaries*, IEEE Trans. Signal Process., **41** (1993), 3397-3415.
- [8] B. K. Natarajan, *Sparse approximate solutions to linear systems*, SIAM J. Comput., **24** (1995), 227-234.
- [9] E. J. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory, **51** (2005), 4203-4215.
- [10] E. J. Candès. *The restricted isometry property and its implications for compressed sensing*, C. R. Math. Acad. Sci. Paris, Serie I, **346** (2008), 589-592.
- [11] S. Foucart and M. J. Lai, *Sparsest solutions of underdetermined linear systems via l_q minimization for $0 < q \leq 1$* , Appl. Comput. Harmon. Anal., **26** (2009), 395-407.
- [12] T. Cai, L. Wang and J. Zhang, *Shifting inequality and recovery of sparse signals*, IEEE Trans. Inform. Theory, **58** (2010), 1300-1308.

- [13] Q. Mo and S. Li, *New bounds on the restricted isometry constant δ_{2k}* , manuscript.
- [14] H. Rauhut, K. Schnass and P. Vandergheynst, *Compressed sensing and redundant dictionaries*, IEEE Trans. Inform. Theory, **54** (2008), 2210-2219.
- [15] D. D. Ge, X. Y. Jiang and Y. Y. Ye, *A note on complexity of l_p minimization*, manuscript.
- [16] R. Chartrand and V. Staneva, *Restricted isometry properties and nonconvex compressive sensing*, Inverse Problems, **24** (2008), 1-14.
- [17] R. Chartrand, *Exact reconstructions of sparse signals via nonconvex minimization*, IEEE Signal Process. Lett., **14** (2007), 707-701.
- [18] R. Saab, R. Chartrand and O. Yilmaz, *Stable sparse approximations via nonconvex optimization*, Proceedings of the 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), (2008), 3885-3888.
- [19] I. Daubechies, R. Devore, M. Fornasier and S. Gunturk, *Iteratively reweighted least squares minimization for sparse recovery*, Comm. Pure. Appl. Math., **13** (2010), 1-38.
- [20] M. E. Davies and R. Gribonval, *Restricted isometry properties where l_p sparse recovery can fail for $0 < p \leq 1$* , IEEE Trans. Inform. Theory, **55** (2010), 2203-2214.
- [21] T. Cai, L. Wang and G. Xu, *New bounds for restricted isometry constants*, IEEE Trans. Inform. Theory, **56** (2010), 4388-4394.
- [22] J. A. Tropp, *Greed is good: Algorithmic results for sparse approximation*, IEEE Trans. Info. Theory, **50** (2004), 2231-2242.
- [23] M. Rudelson and R. Vershynin, *On sparse reconstruction from Fourier and Gaussian measurements*, Comm. Pure Appl. Math., **61** (2008), 1025-1045.
- [24] R. Baraniuk, M. Davenport, R. DeVore and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constr. Approx., **28** (2008), 253-263.
- [25] W. Bajwa, R. Calderbank and S. Jafarpour, *Why Gabor frames? Two fundamental measures of coherence and their geometric significance*, IEEE Trans. Signal Process., (2008), in press.
- [26] M. J. Lai and L. Y. Liu, *A new estimate of restricted isometry constants for sparse solutions*, manuscript.
- [27] D. Needell and J. A. Tropp, *CoSaMP: Iterative signal recovery from noisy samples*, Appl. Comput. Harmon. Anal., **26** (2008), 301-321.